INFORMATION AND COMPUTATION 77, 138–161 (1988)

Synthesizing Inductive Expertise*

DANIEL N. OSHERSON

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

MICHAEL STOB

Calvin College, Grand Rapids, Michigan, 49506

AND

SCOTT WEINSTEIN

University of Pennsylvania, Philadelphia, Pennsylvania 19104

We consider programs that accept descriptions of inductive inference problems and return machines that solve them. Several design specifications for synthesizers of this kind are considered from a recursion-theoretic perspective.

Contents.

Instead of attacking inductive inference problems in piecemeal fashion, one might hope to write a program that synthesizes successful inductive inference machines from input problem descriptions. Such synthesizers might be conceived as accepting information from the user about the nature of the problem to be solved. The present paper examines alternative design specifications for synthesizers of this kind. The specifications vary in terms of

— the nature of the objects to be inferred by synthesized machines, namely, arbitrary r.e. languages versus total recursive functions;

— the quality of problem descriptions input to the synthesizer;

* Preparation of this article was supported by the Office of Naval Research under contract number N001487-K-0401 and a grant from the System Development Foundation. We thank an anonymous reviewer for careful reading of an earlier draft. Correspondence may be addressed to Dan Osherson, E10-034, MIT, Cambridge, MA 02139.

0890-5401/88 $3.00

— the performance characteristics required of synthesized machines (speed, resistance to noisy data, memory limitations, etc.)

For each design specification we ask whether automatic synthesizers of the desired kind exist in principle.

1. Preliminaries

Insofar as possible, notation and terminology are drawn from Osherson, Stob, and Weinstein (1986; henceforth OSW).

The set of natural numbers 0, 1, 2, ... is denoted N. We fix an acceptable indexing $\phi_0, \phi_1, \ldots$ of the partial recursive functions and of their respective domains $W_0, W_1, \ldots$. The class $\{W_i \mid i \in N\}$ is denoted RE. Members of RE are referred to as "languages." We use "L" as a variable over languages. The set $\{\phi_i \mid i \in N\}$ of all partial recursive functions is denoted: F^{rec}. For $i, j \in N$, "$W_{i,j}$" denotes the (finite) set of numbers appearing in the standard enumeration of W_i after j steps of computation. Similarly, $\phi_{i,j}(x)$ is the result of j steps in the computation of $\phi_i(x)$. $X \subseteq N$ is said to be an "index set" for $\{W_j \mid j \in X\}$. We let $D_0, D_1, \ldots$ be the standard enumeration of the finite sets by canonical indices (Rogers, 1967, Section 5.6).

Let $L \in \text{RE}$ be given. A *text for* L is an ω-sequence on L, that is, an infinite listing of all members of L, repetitions allowed, with no members of L^c (the complement of L) in the list. We use "s" and "t" as variables over texts. The set of numbers appearing in text t is denoted: $\text{rng}(t)$.

We let "$\langle \cdot, \cdot \rangle$" code pairs as single integers. The functions $\langle \cdot, \cdot, \cdot \rangle$ etc. are defined from $\langle \cdot, \cdot \rangle$ in the usual fashion. $L \in \text{RE}$ is said to *represent* the set $\{(x, y) \mid \langle x, y \rangle \in L\}$. The class $\{L \in \text{RE} \mid L$ represents a total function$\}$ is denoted: RE_{svt}. ("svt" stands for "single-valued, total.") Thus, RE_{svt} represents the set of graphs of total recursive functions. Whereas it is usual to conceive of inductive inference machines as operating directly on such graphs, it shall here be assumed that graphs are first coded as sets of (single) natural numbers. This will allow uniform treatment of language-inference and function-inference.

Let text t and $n \in N$ be given. The nth member of t (counting from 0) is denoted: $t(n)$. (Thus, the 0th member $t(0)$ of $t = 3, 4, 5, 6, \ldots$ is 3.) The finite initial sequence of length n in t is denoted: $t[n]$. The set $\{t[m] \mid t$ is a text and $m \in N\}$ of all finite sequences in any text is denoted: SEQ. We use "σ," "τ," "δ," as variables over SEQ. The length of $\sigma \in \text{SEQ}$ is denoted: $\text{lh}(\sigma)$. For $\sigma \in \text{SEQ}$ and $m < \text{lh}(\sigma)$, the symbols "$\text{rng}(\sigma)$," "$\sigma(m)$," and "$\sigma[m]$" are interpreted just as for texts. (Notice that $\sigma(\text{lh}(\sigma))$ does not exist, whereas $\sigma[\text{lh}(\sigma)] = \sigma$.) Concatenation among sequences, numbers, and (the beginnings of) texts is denoted $\wedge$.

We assume the existence of a fixed, recursive isomorphism between SEQ and N. Tacit application of this isomorphism allows partial recursive functions to be applied directly to sequences, yielding single natural numbers as outputs.

Let text t, $j \in N$, and $\theta \in F^{\text{rec}}$ be given. θ is *defined on* t just in case for all $m \in N$, $\theta(t[m])$ is defined. θ *converges on* t *to* j just in case (a) θ is defined on t, and (b) for all but finitely many $m \in N$, $\theta(t[m]) = j$. θ *identifies* t just in case θ converges on t to an index for $\text{rng}(t)$. θ *identifies* $L \in \text{RE}$ just in case θ identifies every text for L. θ *identifies* $\mathbf{L} \subseteq \text{RE}$ just in case θ identifies every $L \in \mathbf{L}$. If some $\theta \in F^{\text{rec}}$ identifies $\mathbf{L} \subseteq \text{RE}$, then $\mathbf{L}$ is *identifiable*. Gold (1967) and Blum and Blum (1975) provide interesting examples of both identifiable and nonidentifiable subsets of RE and RE_{svt}. Note that every $\theta \in F^{\text{rec}}$ identifies the empty collection of languages. The foregoing concept of identification corresponds to "EX-identification" in Case and Smith (1983) and elsewhere.

2. Descriptions and Synthesizers

We now provide basic definitions for our study of synthesized inductive inference.

Definition 2A. A mapping $\mathbf{D}: N \to \mathbf{P}(\text{RE})$ (the power-set of RE) is called a *description function*.

Thus, a description function maps each natural number into a collection of languages (as a special case, into a subset of RE_{svt}). In the context of such a function $\mathbf{D}$, a natural number i may be conceived as describing an inductive inference problem, namely, the problem of designing a machine that identifies $\mathbf{D}(i)$. Intuitively, the quality of such a description depends upon the computational transparency of $\mathbf{D}$.

Definition 2B. A partial computable function $S: N \times \text{SEQ} \to N$ is called an (*inductive expertise*) *synthesizer*.

A synthesizer may be conceived as a parameterized inductive inference machine. The parameter is a natural number that codes a collection of languages via a background description function.

Definition 2C. Let $X \subseteq N$, synthesizer S, and description function $\mathbf{D}$ be given.

(i) *S performs X on* $\mathbf{D}$ just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$.

(ii) X is *performable on* $\mathbf{D}$ just in case some synthesizer performs X on $\mathbf{D}$.

Thus, to perform X, S must convert any given problem description $i \in X$ into an inductive inference machine $\lambda\sigma \cdot S(i, \sigma)$ that identifies $\mathbf{D}(i)$. Of course, if $i \in N$ is such that $\mathbf{D}(i)$ is not identifiable, then no $X \subseteq N$ such that $i \in X$ is performable on $\mathbf{D}$.

We now introduce a fundamental description function.

DEFINITION 2D. Let description function $[\cdot]$ be defined as follows. For all $i \in N$, $[i] = \{W_j \mid j \in W_i\}$.

Thus, $[\cdot]$ interprets i as a description of the collection of r.e. sets indexed by W_i. Of course, only r.e. indexable subsets of RE have descriptions under $[\cdot]$.

Let $n \in N$ be an index for N, and let $i_0, i_1, \ldots$ be indices for $D_0, D_1, \ldots$, respectively. Choose $j \in N$ such that $W_j = \{n, i_0, i_1, \ldots\}$. Then, by Gold's theorem (see OSW, Proposition 2.2A), $[j]$ is not identifiable. As a consequence, N is not performable on $[\cdot]$.

Is $\{i \in N \mid [i]$ is identifiable$\}$ performable on $[\cdot]$? To answer this question it is tempting to try to specify total $h \in F^{\text{rec}}$ such that for all $i \in N$, ϕ_i does not identify $W_{h(i)}$. Such an h would be a first step towards defeating any candidate synthesizer for the above set. However, no such function h exists! To see this, suppose otherwise and let total $g \in F^{\text{rec}}$ be such that for all $i \in N$, $\phi_{g(i)}$ is the constant i-function. So, for all $i \in N$, $\phi_{g(i)}$ identifies W_i. By the recursion theorem, let j be a fixed point for $h \circ g$. Then, $\phi_{g(j)}$ identifies $W_j = W_{h(g(j))}$, contradiction. (Compare Case and Smith, 1983, Theorem 2.4.) The foregoing question is answered negatively in the next section as a corollary to a stronger result.

3. PERFORMABILITY IN RE

Our study of synthesized inductive expertise is devoted principally to description functions that return subsets of RE_{svt}. As a prelude, the present section considers a description function that sometimes returns languages in $\text{RE} - \text{RE}_{\text{svt}}$.

To appreciate the content of the following proposition, observe that for all $L, L' \in \text{RE}$, $\{L, L'\}$ is identifiable (see OSW, Exercise 1.4.3C). Hence, for all $i \in N$, if $\text{card}(W_i) = 2$, then $[i]$ is identifiable.

PROPOSITION 3A. *$\{i \in N \mid \text{card}(W_i) = 2\}$ is not performable on* $[\cdot]$.

Thus, no synthesizer can produce successful inductive inference machines from input pairs of arbitrary r.e. indices.

Proof. Let candidate synthesizer S be given. We specify total $h \in F^{\text{rec}}$ such that for all $i \in N$, $\text{card}(W_{h(i)}) = 2$ and $\lambda\sigma \cdot S(i, \sigma)$ does not identify

$\{W_j \mid j \in W_{h(i)}\} = [h(i)]$. Let p be an index for N. Let t^0 be the text: 0, 1, 2, ... Given $i \in N$, we set $W_{h(i)} = \{p, j\}$, where W_j is enumerated as follows.

Stage 0. Find $m_0 \in N$ such that $\text{rng}(t^0[m_0]) \subset W_{S(i, t^0[m_0])}$. If such an m_0 does not exist, then diverge; in this case $W_j = \emptyset$ and $\lambda\sigma \cdot S(i, \sigma)$ fails to identify the text t^0 for $W_p = N$. Otherwise, enumerate $\text{rng}(t^0[m_0])$ into W_j. Set $t^1 = t^0[m_0] \wedge 0 \wedge 0 \wedge \ldots$.

Stage $2n+1$. Find $m_{2n+1} \in N$ such that $m_{2n+1} > m_{2n}$ and $S(i, t^{2n+1}[m_{2n}]) \neq S(i, t^{2n+1}[m_{2n+1}])$. If such an m_{2n+1} does not exist, then diverge; in this case t^{2n+1} is a text for the finite language W_j, and $\lambda\sigma \cdot S(i, \sigma)$ does not identify t^{2n+1} (hence, does not identify W_j). Otherwise, set $t^{2n+2} = t^{2n+1}[m_{2n+1}] \wedge t^0$.

Stage $2n+2$. Find $m_{2n+2} \in N$ such that $m_{2n+2} > m_{2n+1}$ and $\text{rng}(t^{2n+2}[m_{2n+2}]) \subset W_{S(i, t^{2n+2}[m_{2n+2}])}$. If such an m_{2n+2} does not exist, then diverge; in this case $\lambda\sigma \cdot S(i, \sigma)$ fails to identify a text for $W_p = N$. Otherwise, enumerate $\text{rng}(t^{2n+2}[m_{2n+2}])$ into W_j. Set $t^{2(n+1)+1} = t^{2n+2}[m_{2n+2}] \wedge 0 \wedge 0 \wedge \ldots$.

If some even stage of the foregoing construction diverges, then W_j is finite and $\lambda\sigma \cdot S(i, \sigma)$ fails to identify some text for N. If some odd stage diverges, then W_j is finite and $\lambda\sigma \cdot S(i, \sigma)$ fails to identify some text for W_j. If every stage converges, then $W_j = N$ and $\lambda\sigma \cdot S(i, \sigma)$ fails to converge on some text for W_j. Thus, in any case $\lambda\sigma \cdot S(i, \sigma)$ fails to identify $\{W_p, W_j\} = [h(i)]$.

By the recursion theorem, let k be a fixed point for h. Then, $\lambda\sigma \cdot S(k, \sigma)$ fails to identify $[h(k)] = \{W_j \mid j \in W_{h(k)}\} = \{W_j \mid j \in W_k\} = [k]$. Moreover, $\text{card}(W_k) = 2$. ∎

COROLLARY 3A. *$\{i \in N \mid [i]$ is identifiable$\}$ is not performable on $[\cdot]$.*

The proof of Proposition 3A suggests that the nonperformability of $\{i \in N \mid \text{card}(W_i) = 2\}$ on $[\cdot]$ results from the potential equivalence of indices mentioned in a description. This speculation is confirmed by the next proposition. To formulate it, a definition will be helpful.

DEFINITION 3A. $\text{EXT} = \{i \in N \mid (\forall j, k \in W_i)(j \neq k \rightarrow W_j \neq W_k)\}$.

PROPOSITION 3B. *$\{i \in \text{EXT} \mid \text{card}(W_i)$ finite$\}$ is performable on $[\cdot]$.*

Proof. Let i, σ be given. To witness the proposition, $S(i, \sigma)$ is defined as follows. Let $s = \text{lh}(\sigma)$. For each pair $j, k \in W_{i,s}$, define

$$x_{j,k,s} = \mu x \in (W_{j,s} - W_{k,s}) \cup (W_{k,s} - W_{j,s}) \quad \text{if such exists;}$$
$$= \text{undefined, otherwise.}$$

Now define $S(i, \sigma)$ to be the least $j \in W_{i,s}$ such that for all $k \in W_{i,s}$, if $x_{j,k,s}$ is defined, then $x_{j,k,s} \in \text{rng}(\sigma)$ if and only if $x_{j,k,s} \in W_{j,s}$. (And define $S(i, \sigma) = 0$ if no such j exists.) To see that $S(i, \sigma)$ identifies $[i]$ if $i \in \text{EXT}$ and W_i is finite, let s_0 be such that $W_i = W_{i,s_0}$. Suppose also that for all $s \geqslant s_0$, each $x_{j,k,s}$ is defined and $x_{j,k,s} = \mu x \in (W_j - W_k) \cup (W_k - W_j)$. Now let t be any text for W_j, $j \in W_i$. Let $s_1 \geqslant s_0$ be such that for all $k \in W_i$, $x_{j,k,s} \in W_j$ implies $x_{j,k,s_1} \in t[s_1]$. Then for all $s \geqslant s_1$, $S(i, \sigma) = j$ and $S(i, \sigma)$ identifies t. ∎

The finiteness qualification in the preceding proposition is essential. This is shown by the following result.

PROPOSITION 3C. *$\{i \in \text{EXT} \mid [i]$ is identifiable$\}$ is not performable on* $[\cdot]$.

Proposition 3C is proved in Section 4.

4. PERFORMABILITY IN RE_{svt}

We now introduce a description function that returns subsets of RE_{svt}. Performability on this function will be our principal concern throughout the remainder of this paper.

DEFINITION 4A. Let description function $[\cdot]_{\text{svt}}$ be defined as follows. For all $i \in N$, $[i]_{\text{svt}} = [i] \cap \text{RE}_{\text{svt}}$.

Thus, for all $i \in N$, $[i]_{\text{svt}} = \{W_j \in \text{RE}_{\text{svt}} \mid j \in W_i\}$. In the context of $[\cdot]_{\text{svt}}$ an index i may be conceived as an "impure" index set for some subset of RE_{svt}—"impure" in the sense of possibly holding irrelevant indices for languages outside of RE_{svt}. The presence of such irrelevant indices is central to performability on $[\cdot]_{\text{svt}}$. To see this, let n be an index for N. It follows from Proposition 4.2.1B of OSW (due to Gold, 1967) that $[n]_{\text{svt}}$ $(= \text{RE}_{\text{svt}})$ is not identifiable. In contrast, if indices for members of $\text{RE} - \text{RE}_{\text{svt}}$ are not present in descriptions, then performability on $[\cdot]_{\text{svt}}$ is always possible. This is the content of the next proposition. A definition will aid its formulation.

DEFINITION 4B. $\text{SVT} = \{i \in N \mid (\forall j \in W_i)\ (W_j \in \text{RE}_{\text{svt}})\}$.

PROPOSITION 4A. SVT *is performable on* $[\cdot]_{\text{svt}}$.

Proof. We exhibit synthesizer S that performs SVT on $[\cdot]_{\text{svt}}$. Given $i \in N$, $\sigma \in \text{SEQ}$, define $S(i, \sigma)$ to be the first j to appear in the standard enumeration of W_i such that $\mathrm{rng}(\sigma) \subseteq W_j$. It is clear that if every index in W_i is for a member of $\mathrm{RE}_{\mathrm{svt}}$, then $\lambda\sigma \cdot S(i, \sigma)$ identifies $[i] = [i]_{\mathrm{svt}}$. ∎

Proposition 4A leads us to ask whether all communication to synthesizers can be carried out under $[\cdot]_{\mathrm{svt}}$ through indices in SVT. Put differently, we ask whether every identifiable subset of $\mathrm{RE}_{\mathrm{svt}}$ is included in some member of $\{[i]_{\mathrm{svt}} \mid i \in \mathrm{SVT}\}$. The next proposition answers this question negatively.

DEFINITION 4C. $\theta \in F^{\mathrm{rec}}$ is *self-indexing* just in case the least $x \in N$ such that $\theta(x) = 1$ is an index for θ. The collection $\{L \in \mathrm{RE}_{\mathrm{svt}} \mid L$ represents a self-indexing function$\}$ is denoted: $\mathrm{RE}_{\mathrm{si}}$.

PROPOSITION 4B. *$\mathrm{RE}_{\mathrm{si}}$ is identifiable and for every $i \in \mathrm{SVT}$, $\mathrm{RE}_{\mathrm{si}} \not\subseteq [i]_{\mathrm{svt}}$.*

Proof. The fact that $\mathrm{RE}_{\mathrm{si}}$ is identifiable is due to Blum and Blum (1975). (The obvious method works.) To show that for all $i \in \mathrm{SVT}$, $\mathrm{RE}_{\mathrm{si}} \not\subseteq [i]_{\mathrm{svt}}$, let total recursive functions p and q be such that:

(a) for all $k \in N$, $W_{q(k)}$ represents ϕ_k; and

(b) for all $k \in N$, if $W_k \in \mathrm{RE}_{\mathrm{svt}}$, then W_k represents $\phi_{p(k)}$.

Let i be given. If $W_i = \varnothing$, then $\mathrm{RE}_{\mathrm{si}} \not\subseteq [i]_{\mathrm{svt}}$ (it will be seen below that $\mathrm{RE}_{\mathrm{si}} \neq \varnothing$). Assume that $W_i \neq \varnothing$. We specify total recursive function h such that for all $x \in N$:

(i) $\phi_{h(x)}(y) = 0$ for all $y < x$;

(ii) $\phi_{h(x)}(x) = 1$;

(iii) $\phi_{h(x)}$ is total; and

(iv) $W_{q(h(x))} \notin [i]_{\mathrm{svt}}$.

Application of the recursion theorem then yields $k \in N$ such that $\phi_{h(k)} = \phi_k$, so:

(i′) $\phi_k(y) = 0$ for all $y < k$;

(ii′) $\phi_k(k) = 1$;

(iii′) ϕ_k is total; and

(iv′) $W_{q(k)} \notin [i]_{\mathrm{svt}}$.

Hence, ϕ_k is self-indexing and total, and $W_{q(k)}$ witnesses that $\mathrm{RE}_{\mathrm{si}} \not\subseteq [i]_{\mathrm{svt}}$.

It remains to specify h. Let total recursive g be such that $\mathrm{range}(g) = \{p(j) \mid j \in W_i\}$. Then $\phi_{g(z)}$ is total for all $z \in N$. Given $x \in N$, let $h(x)$ be a uniformly effectively constructed index for the function θ such that for all $y \in N$,

$$\begin{aligned}\theta(y) &= 0 && \text{if } y < x;\\ &= 1 && \text{if } y = x;\\ &= \phi_{g(y-(x+1))}(y) + 1 && \text{if } y > x.\end{aligned}$$

It is easy to verify that for all $x \in N$, $\phi_{h(x)}$ meets conditions (i)–(iv). ▮

On the other hand, $[\cdot]_{\text{svt}}$ does suffice to communicate every identifiable subset of RE_{svt}, in some cases via "impure" descriptions. This is the content of the next proposition.

PROPOSITION 4C. *If* $\mathbf{L} \subseteq \text{RE}_{\text{svt}}$ *is identifiable, then there is* $i \in N$ *such that* $\mathbf{L} \subseteq [i]_{\text{svt}}$ *and* $[i]_{\text{svt}}$ *is identifiable.*

Proof. The proposition is a corollary of OSW, Lemma 4.3.4B (due to Mark Fulk). ▮

Remaining with the general case (in which $i \notin \text{SVT}$ is possible), we now consider whether performability is assured by limiting attention to indices i such that $[i]_{\text{svt}}$ is identifiable. The next proposition provides a negative answer to this question.

PROPOSITION 4D. $\{i \in N \mid [i]_{\text{svt}}$ *is identifiable*$\}$ *is not performable on* $[\cdot]_{\text{svt}}$.

Proof. Let candidate synthesizer S be given. We exhibit total $h \in F^{\text{rec}}$ such that for all $i \in N$, $\lambda\sigma \cdot S(i, \sigma)$ does not identify $\{W_j \in \text{RE}_{\text{svt}} \mid j \in W_{h(i)}\} = [h(i)]_{\text{svt}}$. Moreover, for all $i \in N$, $[h(i)]_{\text{svt}}$ will be seen to be identifiable. The recursion theorem then yields $k \in N$ such that $\lambda\sigma \cdot S(k, \sigma)$ does not identify $[h(k)]_{\text{svt}} = [k]_{\text{svt}}$ and $[k]_{\text{svt}}$ is identifiable.

Given $i \in N$, $W_{h(i)}$ is enumerated in stages. Let t^0 be the text: $\langle 0, 0\rangle$, $\langle 1, 0\rangle$, $\langle 2, 0\rangle$, ... Let p_0 be an index for $\text{rng}(t^0)$ and enumerate p_0 into $W_{h(i)}$.

Stage 0. Find $m_0, q_0 \in N$ such that $q_0 > m_0$ and $\langle q_0, 0\rangle \in W_{S(i,t[m_0])}$. If no such m_0, q_0 exist, then $\lambda\sigma \cdot S(i, \sigma)$ fails to identify a text for W_{p_0}. Otherwise, set $t^1 = t^0[m_0] \wedge \langle m_0, 1\rangle \wedge \langle m_0 + 1, 1\rangle \wedge \ldots$. Let p_1 be an index for $\text{rng}(t^1)$, and enumerate p_1 into $W_{h(i)}$.

Stage $2n+1$. Find $m_{2n+1} > m_{2n}$ such that $S(i, t^{2n+1}[m_{2n+1}]) \neq S(i, t^{2n}[m_{2n}])$. If no such m_{2n+1} exists, then $\lambda\sigma \cdot S(i, \sigma)$ fails to identify a text for $W_{p_{2n+1}}$. Otherwise, set $t^{2n+2} = t^{2n+1}[m_{2n+1}] \wedge \langle m_{2n+1}, 0\rangle \wedge \langle m_{2n+1} + 1, 0\rangle \wedge \ldots$. Let p_{2n+2} be an index for $\text{rng}(t^{2n+2})$ and enumerate p_{2n+2} into $W_{h(i)}$.

Stage $2n+2$. Find m_{2n+2}, $q_{2n+2} \in N$ such that $q_{2n+2} > m_{2n+2}$ and $\langle q_{2n+2}, 0\rangle \in W_{S(i,\, t^{2n+2}[m_{2n+2}])}$. If no such m_{2n+2}, q_{2n+2} exist, then $\lambda\sigma \cdot S(i, \sigma)$ fails to identify a text for $W_{p_{2n+2}}$. Otherwise, set $t^{2(n+1)+1} = t^{2n+2}[m_{2n+2}] \wedge \langle m_{2n+2}, 1\rangle \wedge \langle m_{2n+2}+1, 1\rangle \wedge \ldots$. Let $p_{2(n+1)+1}$ be an index for $\text{rng}(t^{2(n+1)+1})$ and enumerate $p_{2(n+1)+1}$ into $W_{h(i)}$.

Finally, we define a "diagonal" language $L \in \text{RE}$ as follows. For all $n \in N$, if stage n in the foregoing construction terminates, enumerate $\text{rng}(t^n[m_n])$ into L. Let d be an index for L and enumerate d into $W_{h(i)}$. (Since the construction is effective, d can be enumerated into $W_{h(i)}$ just after p_0.)

If some stage n in the foregoing construction diverges, then $\lambda\sigma \cdot S(i, \sigma)$ converges to an incorrect index on some text for $W_{p_n} \in \text{RE}_{\text{svt}}$. In this case W_d is finite (and hence not a member of RE_{svt}). On the other hand, if all stages in the construction terminate, then $W_d \in \text{RE}_{\text{svt}}$ and $\lambda\sigma \cdot S(i, \sigma)$ fails to converge on some text for W_d (namely, the text obtained from the sequences $t^n[m_n]$ in the obvious way). Thus, in either case, $\lambda\sigma \cdot S(i, \sigma)$ fails to identify $[h(i)] \cap \text{RE}_{\text{svt}} = [h(i)]_{\text{svt}}$. Moreover, it is easy to verify that for all $i \in N$, $[h(i)]_{\text{svt}}$ is identifiable. ∎

The preceding construction is easily adapted for the proof of Proposition 3C.

Proof of Proposition 3C. Let h be as defined in the proof of Proposition 4D, and let $i \in N$ be given. Since $\lambda\sigma \cdot S(i, \sigma)$ does not identify $[h(i)]_{\text{svt}}$, $\lambda\sigma \cdot S(i, \sigma)$ does not identify $[h(i)]$. It thus suffices to observe that $[h(i)]$ is identifiable and $h(i) \in \text{EXT}$. ∎

In contrast to Proposition 4D, the next result shows that the set of indices for finite sets can be performed on $[\cdot]_{\text{svt}}$. (Compare Case and Smith, 1983, Theorem 2.9.)

PROPOSITION 4E. $\{i \in N \mid \text{card}(W_i)$ *finite*$\}$ *is performable on* $[\cdot]_{\text{svt}}$.

Proof. We say that $D, E \subseteq N$ *conflict* just in case there is $n \in N$ such that for some $x, y \in N$, $x \neq y$, $\langle n, x\rangle \in D$, and $\langle n, y\rangle \in E$. By Rogers (1967, Theorem XVI, Chapt. 5) there is a recursive function f such that $W_{f(i)}$ is single-valued and if W_i is single-valued then $W_{f(i)} = W_i$.

The following synthesizer S witnesses the proposition. Given $i \in N$, $\sigma \in \text{SEQ}$, S computes $D = \{j \in W_{i,\text{lh}(\sigma)} \mid W_{j,\text{lh}(\sigma)}$ and $\text{rng}(\sigma)$ do not conflict$\}$. S then computes an index k for $\bigcup\{W_j \mid j \in D\}$ and emits $f(k)$. ∎

The next six sections are devoted to special properties that one might wish to design into synthesizers.

5. Direct Synthesis

Let $i \in N$ be given. Then, for every $L \in [i]_{svt}$ there is $j \in W_i$ such that $L = W_j$. This suggests that synthesizers working under $[\cdot]_{svt}$ be designed to limit their conjectures to indices in W_i when parameterized by i. Would such a design affect performability?

DEFINITION 5A. Let $i \in N$, $\theta \in F^{rec}$, and description function $\mathbf{D}$ be given.

(i) θ *identifies* $\mathbf{D}(i)$ *from* i just in case (a) θ identifies $\mathbf{D}(i)$, and (b) for all $\sigma \in \mathrm{SEQ}$, $\theta(\sigma) \in W_i$.

(ii) $\mathbf{D}(i)$ is *identifiable from* i just in case some $\theta \in F^{rec}$ identifies $\mathbf{D}(i)$ from i.

Evidently, $\mathbf{D}(i)$ is identifiable from i only if $\mathbf{D}(i)$ is included in $[i]$. Observe as well that if θ identifies $\mathbf{D}(i) \neq \varnothing$ from i, then θ is total. We illustrate the definition with the following result.

PROPOSITION 5A. *For all* $i \in \mathrm{SVT}$, $[i]_{svt}$ *is identifiable from* i.

Proof. $\theta \in F^{rec}$ may identify $[i]_{svt}$ in "induction-by-enumeration" fashion, as in the proof of Proposition 4A. ▮

DEFINITION 5B. Let $X \subseteq N$, synthesizer S, and description function $\boldsymbol{D}$ be given.

(i) *S performs X on D directly* just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$ from i.

(ii) *X* is *performable on* $\mathbf{D}$ *directly* just in case some synthesizer performs X on $\mathbf{D}$ directly.

PROPOSITION 5B. SVT *is performable on* $[\cdot]_{svt}$ *directly.*

Proof. The synthesizer S in the proof of Proposition 4A witnesses the proposition. ▮

We now consider direct performability outside of SVT. The following proposition shows that the directness requirement can obstruct performability on $[\cdot]_{svt}$, even if attention is limited to indices j that (taken individually) permit identification of $[j]_{svt}$ from j.

PROPOSITION 5C. *There is a set* $X \subseteq N$ *such that*

(i) *X is performable on* $[\cdot]_{svt}$;

(ii) *for all* $j \in X$, $[j]_{svt}$ *is identifiable from* j; *but*

(iii) *X is not performable on* $[\cdot]_{svt}$ *directly.*

Proof. Let total $h \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{h(i)}$ contains all numbers of either of the following forms:

(a) $\langle x, 0 \rangle$, where $W_{i,x} \subset W_{i,x+1}$;

(b) $\langle x, 1 \rangle$, where $W_{i,x} = W_{i,x+1}$ but for some $k \in N$, $W_{i,x} \subset W_{i,x+k}$.

Observe that for all $i \in N$, $W_{h(i)} \in \text{RE}_{\text{svt}}$ iff W_i is infinite.

For $i \in N$, let text $t^i = \langle 0, n_0 \rangle, \langle 1, n_1 \rangle, \ldots$, where for all $x \in N$, $n_x = 0$ if $W_{i,x} \subset W_{i,x+1}$ and $n_x = 1$ otherwise. Observe that for all $i \in N$:

(a) $\text{rng}(t^i) \in \text{RE}_{\text{svt}}$; and

(b) $\text{rng}(t^i) = W_{h(i)}$ iff W_i is infinite (otherwise, $\text{rng}(t^i)$ properly includes $W_{h(i)}$).

Recall that $D_0, D_1, \ldots$ is the standard enumeration of the finite sets. Let total $g \in F^{\text{rec}}$ be such that for all $m \in N$,

$$W_{g(m)} = \{\langle x, n_x \rangle \mid n_x = 0 \text{ if } x \in D_m \text{ and } n_x = 1 \text{ otherwise}\}.$$

Observe that for all $i \in N$, if W_i is infinite then $\text{rng}(t^i) \neq W_{g(m)}$ for all $m \in N$; whereas if W_i is finite, then $\text{rng}(t^i) = W_{g(m)}$ for some $m \in N$.

Let total, one–one $d \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{d(i)} = \{h(i)\} \cup \{g(m) \mid m \in N\}$. ($d$ can be chosen to be one–one by appropriate "padding.") The set X of the proposition is taken to be $\{d(i) \mid i \in N\}$.

To verify clause (i) of the proposition, let synthesizer S operate as follows. On input $(j, \sigma) \in N \times \text{SEQ}$, S relies on the one–one nature of d to find $i \in N$ such that $j = d(i)$ (if no such i is found, $S(j, \sigma)$ diverges). S then computes an index z for $\text{rng}(t^i)$. (Of course, z need not be a member of $W_{d(i)}$.) Next, S conjectures the first index u in the list z, $g(0)$, $g(1)$, ... such that $\text{rng}(\sigma) \subseteq W_u$. Since the list holds only indices for languages in RE_{svt}, $\lambda\sigma \cdot S(d(i), \sigma)$ identifies $\{W_z\} \cup \{W_{g(m)} \mid m \in N\}$. In particular, if W_i is infinite, then $\lambda\sigma \cdot S(d(i), \sigma)$ identifies $W_{h(i)} = W_z$. If W_i is finite, then $W_{h(i)} \notin \text{RE}_{\text{svt}}$ so it need not be identified. It is thus clear that S performs X.

To verify clause (ii) of the proposition, let $i \in N$ be given. Let $\theta_1 \in F^{\text{rec}}$ operate as follows. Given $\sigma \in \text{SEQ}$, θ_1 conjectures the first index u in the list $h(i)$, $g(0)$, $g(1)$, ... such that $\text{rng}(\sigma) \subseteq W_u$. It is easy to see that if W_i is infinite, then θ_1 identifies $[d(i)]_{\text{svt}}$ from $d(i)$. On the other hand, let $\theta_2 \in F^{\text{rec}}$ operate as follows. Given $\sigma \in \text{SEQ}$, θ_2 conjectures the first index u in the list $g(0)$, $g(1)$, ... such that $\text{rng}(\sigma) \subseteq W_u$. It is easy to see that if W_i is finite, then θ_2 identifies $[d(i)]_{\text{svt}}$ from $d(i)$. So, one of θ_1, θ_2 witnesses the identifiability of $[d(i)]_{\text{svt}}$ from $d(i)$.

Finally, to verify clause (iii) of the proposition, suppose that synthesizer S performs X on $[\cdot]_{\text{svt}}$ directly. Then, for all $i \in N$, $\lambda\sigma \cdot S(d(i), \sigma)$ is total. If W_i is infinite, then t^i is a text for $W_{h(i)}$. So $\lambda\sigma \cdot S(d(i), \sigma)$ identifies t^i. Moreover, for all $k \in W_{d(i)}$, $W_k = W_{h(i)}$ iff $k = h(i)$, as easily seen. Therefore, by the directness of S, $\lambda\sigma \cdot S(d(i), \sigma)$ converges on t^i to $h(i)$. In summary, for all $i \in N$, if W_i is infinite then:

$$(\exists r \in N)(\forall u \in N)(u > r \rightarrow S(d(i), t^i[u]) = h(i)), \tag{1}$$

where $t^i[u]$ can be recovered effectively from i, u and where h, d, and $\lambda\sigma \cdot S(d(i), \sigma)$ are total recursive.

If W_i is finite then there is $m \in N$ such that t^i is a text for $W_{g(m)}$. Moreover, for all $k \in W_{d(i)}$, $W_k = W_{g(m)}$ iff $k = g(m)$. Therefore, by the directness of S, for some $m \in N$, $\lambda\sigma \cdot S(d(i), \sigma)$ converges on t^i to $g(m) \neq h(i)$. In summary, for all $i \in N$, if W_i is finite then:

$$(\forall r \in N)\ (\exists u \in N)\ (u > r \text{ and } S(d(i), t^i[u]) \neq h(i)). \tag{2}$$

Since (2) is the negation of (1), the set $\{i \mid W_i \text{ infinite}\}$ is thus exhibited as Σ_2, contradicting its Π_2 completeness (see Rogers, 1967, Section 14.8). ∎

As a corollary to the proof, note that the set X of the proposition may be chosen to be r.e. Of independent interest is the following result, which will be proved by modifying the foregoing proof (compare Proposition 5A).

PROPOSITION 5D. *There is $j \in N$ such that $[j]_{\text{svt}}$ is identifiable, but $[j]_{\text{svt}}$ is not identifiable from j.*

Proof. Let total $h \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{h(i)}$ contains all numbers of either of the following forms:

(a) $\langle x, \langle 0, i \rangle \rangle$, where $W_{i,x} \subset W_{i,x+1}$;

(b) $\langle x, \langle 1, i \rangle \rangle$, where $W_{i,x} = W_{i,x+1}$ but for some $k \in N$, $W_{i,x} \subset W_{i,x+k}$.

For $i \in N$, let text $t^i = \langle 0, n_0 \rangle, \langle 1, n_1 \rangle, \ldots$, where for all $x \in N$, $n_x = \langle 0, i \rangle$ if $W_{i,x} \subset W_{i,x+1}$ and $n_x = \langle 1, i \rangle$, otherwise. Let total $g \in F^{\text{rec}}$ be such that for all $i, m \in N$,

$$W_{g(\langle i,m \rangle)} = \{\langle x, n_x \rangle \mid n_x = \langle 0, i \rangle \text{ if } x \in D_m \text{ and } n_x = \langle 1, i \rangle \text{ otherwise}\}.$$

Let total $f \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{f(i)} = \{h(i)\} \cup \{g(\langle i, m \rangle \mid m \in N\}$. Finally, choose $j \in N$ such that $W_j = \bigcup_{i \in N} W_{f(i)}$.

From this point the proof of Proposition 5D parallels that given for clauses (i) and (iii) of Proposition 5C. ∎

6. Text-Efficient Synthesis

We would like synthesized inductive inference machines to arrive at correct hypotheses as fast as possible. In particular, such machines should examine a minimum number of data before beginning to converge. The present section considers the prospects for synthesizing text-efficient inductive inference machines. Some preliminary definitions are necessary. Let **T** be the class of texts for languages in RE. In what follows, "IP" may be read as "identification point."

Definition 6A. Total function $\text{IP}: F^{\text{rec}} \to N \cup \{\boldsymbol{\omega}\}$ is defined as follows. For all $\theta \in F^{\text{rec}}$, $t \in \mathbf{T}$:

$$\begin{aligned}\text{IP}(\theta, t) &= \boldsymbol{\omega}, \qquad \text{if } \theta \text{ does not identify } t;\\ &= \mu n[\theta(t[n]) = \theta(t[n+k]) \text{ for all } k \in N],\\ &\quad \text{otherwise.}\end{aligned}$$

The next definition construes text-efficiency in terms of identification points.

Definition 6B (Gold, 1967). Let $\theta, \theta' \in F^{\text{rec}}$ and $\mathbf{L} \subseteq \text{RE}$ be given.

(i) θ *identifies* **L** *strictly faster than* θ' just in case (a) $\text{IP}(\theta, t) \leqslant \text{IP}(\theta', t)$ for all texts t for all $L \in \mathbf{L}$, and (b) $\text{IP}(\theta, s) < \text{IP}(\theta', s)$ for some text s for some $L \in \mathbf{L}$.

(ii) θ *identifies* **L** *text-efficiently* just in case (a) θ identifies **L**, and (b) no $\theta' \in F^{\text{rec}}$ identifies **L** strictly faster than θ.

(iii) **L** is *identifiable text-efficiently* just in case some $\theta \in F^{\text{rec}}$ identifies **L** text-efficiently.

We illustrate the definition with the following result.

Proposition 6A. (i) *For all* $i \in \text{SVT}$, $[i]_{\text{svt}}$ *is identifiable text-efficiently.*

(ii) *There is* $i \in N$ *such that* $[i]_{\text{svt}}$ *is not identifiable text-efficiently.*

Proof. (i) is an immediate consequence of Proposition 8.2.4B of OSW. (ii) may be derived by an adaptation of the proof of Proposition 6C, below, in analogous fashion to the proof of Proposition 5D. We omit the details. ∎

Text-efficient synthesizers may now be defined on the basis of Definition 6B.

DEFINITION 6C. Let $X \subseteq N$, synthesizer S, and description function $\mathbf{D}$ be given.

(i) *S performs X on* $\mathbf{D}$ *text-efficiently* just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$ text-efficiently.

(ii) *X is performable on* $\mathbf{D}$ *text-efficiently* just in case some synthesizer performs X on $\mathbf{D}$ text-efficiently.

PROPOSITION 6B. SVT *is performable on* $[\cdot]_{\text{svt}}$ *text-efficiently.*

Proof. Let synthesizer S be as defined in the proof of Proposition 4A. Proposition 8.2.4A(ii) of OSW (due to Gold, 1967) implies that for all $i \in N$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $[i]_{\text{svt}}$ text-efficiently. ▮

We now consider text-efficient performability in the more general context (not restricted to SVT). The next proposition shows that the requirement of text-efficiency can obstruct performability on $[\cdot]_{\text{svt}}$, even if attention is limited to indices j that (taken individually) permit identification of $[j]_{\text{svt}}$ text-efficiently.

PROPOSITION 6C. *There is a set $X \subseteq N$ such that*

(i) *X is performable on* $[\cdot]_{\text{svt}}$;

(ii) *for all $j \in X$, $[j]_{\text{svt}}$ is identifiable text-efficiently*; *but*

(iii) *X is not performable on* $[\cdot]_{\text{svt}}$ *text-efficiently.*

Proof. Let total $h \in F^{\text{rec}}$ be such that for all $i \in N$:

(a) $\{\langle 0, 0\rangle, \langle 1, 8\rangle\} \subseteq W_{h(i)}$; and

(b) $W_{h(i)} \in \text{RE}_{\text{svt}}$ iff $W_i = \varnothing$.

Let total $g \in F^{\text{rec}}$ be such that for all $i \in N$:

(a) $\{\langle 0, 0\rangle, \langle 1, 9\rangle\} \subseteq W_{g(i)}$; and

(b) $W_{g(i)} \in \text{RE}_{\text{svt}}$ iff $W_i \neq \varnothing$.

We leave it to the reader to verify that such h, g exist. Let total, one–one $d \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{d(i)} = \{h(i), g(i)\}$. We choose X of the proposition to be $\{d(i) \mid i \in N\}$.

Clause (i) follows from Proposition 3B, and (ii) is easy to verify. As for (iii), suppose that synthesizer S performs X text-efficiently. Then, for all $i \in N$, $S(d(i), \langle 0, 0\rangle)$ is defined. Let $i \in N$ be such that $W_i = \varnothing$. Then $W_{h(i)} \in \text{RE}_{\text{svt}}$ and $W_{g(i)} \notin \text{RE}_{\text{svt}}$. So, $S(d(i), \langle 0, 0\rangle)$ is an index for $W_{h(i)}$. For otherwise it is easy to specify $\theta \in F^{\text{rec}}$ that is strictly faster than $\lambda\sigma \cdot S(i, \sigma)$ on $[d(i)]_{\text{svt}}$. Such a θ need merely converge immediately to $h(i)$ on every text. Similarly, if $W_i \neq \varnothing$, then $S(d(i), \langle 0, 0\rangle)$ is an index for

$W_{g(i)}$. It follows that $W_i = \emptyset$ if and only if $\langle 1, 8 \rangle \in W_{S(d(i), \langle 0,0 \rangle)}$, where $\lambda i \cdot S(d(i), \langle 0, 0 \rangle)$ is total. This contradicts the fact that $\{i \mid W_i = \emptyset\}$ is not r.e. ∎

As a corollary to the foregoing proof, note that the set X of Proposition 6C can be chosen so that X is r.e. and for all $j \in X$, $\text{card}(W_j) = 2$.

7. Consistency

Consistent synthesizers produce inductive inference machines whose conjectures "cover" the data that provoke them. This is made precise as follows.

DEFINITION 7A (Angluin, 1980). (i) $\theta \in F^{\text{rec}}$ is *consistent* just in case for all $\sigma \in \text{SEQ}$, $\text{rng}(\sigma) \subseteq W_{\theta(\sigma)}$.

(ii) $\mathbf{L} \subseteq \text{RE}$ is *identifiable consistently* just in case some consistent $\theta \in F^{\text{rec}}$ identifies $\mathbf{L}$.

Note that consistent functions are total.

DEFINITION 7B. Let $X \subseteq N$, synthesizer S, and description function $\mathbf{D}$ be given.

(i) *S performs X on* $\mathbf{D}$ *consistently* just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$ and $\lambda\sigma \cdot S(i, \sigma)$ is consistent.

(ii) X is *performable on* $\mathbf{D}$ *consistently* just in case some synthesizer performs X on $\mathbf{D}$ consistently.

PROPOSITION 7A. SVT *is performable on* $[\cdot]_{\text{svt}}$ *consistently.*

Proof. The following synthesizer S witnesses the proposition. Given $(i, \sigma) \in N \times \text{SEQ}$, S enumerates $W_{i,\text{lh}(\sigma)}$. If no $j \in W_{i,\text{lh}(\sigma)}$ is such that $\text{rng}(\sigma) \subseteq W_j$, then S conjectures an index for $\text{rng}(\sigma)$; otherwise, S conjectures the least $j \in W_{i,\text{lh}(\sigma)}$ with this property. If $i \in \text{SVT}$, then $\lambda\sigma \cdot S(i, \sigma)$ is consistent since in this case every $j \in W_i$ is such that $\text{rng}(\sigma) \subseteq W_j$ is confirmed or disconfirmed after a finite computation. On the other hand, if $i \notin \text{SVT}$, then $\lambda\sigma \cdot S(i, \sigma)$ is not required to be consistent or even total. Finally, it is clear that $\lambda\sigma \cdot S(i, \sigma)$ identifies $[\cdot]_{\text{svt}}$ for every $i \in \text{SVT}$. ∎

Returning to the more general context, we now show that the consistency requirement can obstruct performability on $[\cdot]_{\text{svt}}$ even if attention is limited to indices j that (taken individually) permit identification of $[j]_{\text{svt}}$ consistently.

PROPOSITION 7B. *There is a set* $X \subseteq N$ *such that*

(i) X *is performable on* $[\cdot]_{\text{svt}}$;

(ii) *for all* $j \in X$, $[j]_{\text{svt}}$ *is identifiable consistently*; *but*

(iii) X *is not performable on* $[\cdot]_{\text{svt}}$ *consistently.*

We rely upon the following lemma, given as Proposition 2.1A in OSW.

LEMMA 7A (Blum and Blum, 1975). *Let* $\theta \in F^{\text{rec}}$ *identify* $L \in \text{RE}$. *Then there is* $\sigma \in \text{SEQ}$ *such that* (a) $\text{rng}(\sigma) \subseteq L$, (b) $W_{\theta(\sigma)} = L$, *and* (c) *for all* $\tau \in \text{SEQ}$, *if* $\text{rng}(\tau) \subseteq L$, *then* $\theta(\sigma \wedge \tau) = \theta(\sigma)$.

Proof of Proposition 7B. Let total $h \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{h(i)} = \{\langle x, y\rangle \mid \phi_i(x) = y\}$. Let total, one–one $d \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{d(i)} = \{h(i)\}$. Then, for all $i \in N$, $[d(i)]_{\text{svt}} = \{W_{h(i)}\}$, if ϕ_i is total, and is empty otherwise. Let X of the proposition be $\{d(i) \mid i \in N\}$.

Clause (i) of the proposition follows from Proposition 3B, and (ii) is easy to verify. For a contradiction, suppose that synthesizer S performs X consistently. Then:

(1) For all $i \in N$, $\lambda\sigma \cdot S(d(i), \sigma)$ is constent (and hence total).

(2) For all $i \in N$, if ϕ_i is total, then $\lambda\sigma \cdot S(d(i), \sigma)$ identifies $\{\langle x, y\rangle \mid \phi_i(x) = y\}$.

From (2) and Lemma 7A:

(3) For all $i \in N$, if ϕ_i is total, then there is $\sigma \in \text{SEQ}$ such that for all $x, y \in N$, if $\phi_i(x) = y$ then $S(d(i), \sigma) = S(d(i), \sigma \wedge \langle x,y\rangle)$.

By (3) and (1):

(4) For all $i \in N$, if ϕ_i is total, then there is $\sigma \in \text{SEQ}$ such that for all $x, y \in N$, $\phi_i(x) = y$ iff $S(d(i), \sigma) = S(d(i), \sigma \wedge \langle x, y\rangle)$, where for all $j \in N$, $\lambda\sigma \cdot S(d(j), \sigma)$ is total.

However, (4) is impossible, as revealed by the following diagonalization. Let SEQ be indexed as $\sigma^0, \sigma^1, \ldots$. Define total $f \in F^{\text{rec}}$ as follows. For all $x, y \in N$,

$$f(\langle x, y\rangle) = 0 \quad \text{if} \quad S(d(x), \sigma^y) \neq S(d(x), \sigma^y \wedge \langle\langle x, y\rangle, 0\rangle)$$
$$= 1 \quad \text{otherwise.}$$

Let z be an index for f. Then, if σ^w has the properties given in (4), we have $S(d(z), \sigma^w) \neq S(d(z), \sigma^w \wedge \langle\langle z, w\rangle, 0\rangle)$ iff $f(\langle z, w\rangle) = 0$ iff $S(d(z), \sigma^w) = S(d(z), \sigma^w \wedge \langle\langle z, w\rangle, 0\rangle$. ▮

As a corollary to the foregoing proof, note that the set X of

Proposition 7B can be chosen so that X is r.e. and for all $j \in X$, $\text{card}(W_j) = 1$.

8. Reliability

Inductive inference machines that invariably signal the incorrectness of prior, false conjectures are called "reliable." We would like our synthesizers to produce reliable machines.

Definition 8A (Minicozzi, cited in Blum and Blum, 1975). (i) $\theta \in F^{\text{rec}}$ is *reliable* just in case (a) θ is total, and (b) for all texts t for any $L \in \text{RE}_{\text{svt}}$, if θ converges on t then θ identifies t.

(ii) $\mathbf{L} \subseteq \text{RE}_{\text{svt}}$ is *identifiable reliably* just in case some reliable $\theta \in F^{\text{rec}}$ identifies $\mathbf{L}$.

Thus, reliable machines never converge to an incorrect index on any text for a language in RE_{svt}.

Definition 8B. Let $X \subseteq N$, synthesizer S, and description function $\mathbf{D}$ be given.

(i) *S performs X on $\mathbf{D}$ reliably* just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$ and $\lambda\sigma \cdot S(i, \sigma)$ is reliable.

(ii) *X is performable on $\mathbf{D}$ reliably* just in case some synthesizer performs X on $\mathbf{D}$ reliably.

Proposition 8A. *SVT is performable on $[\cdot]_{\text{svt}}$ reliably.*

Proof. With a slight modification, the synthesizer S described in the proof of Proposition 7A witnesses the present proposition. The modification is that S must conjecture $\text{lh}(\sigma)$ in case no $j \in W_{i,\text{lh}(\sigma)}$ is such that $\text{rng}(\sigma) \subseteq W_j$. ∎

Returning to the general context, the next proposition shows that the reliability requirement can obstruct performability on $[\cdot]_{\text{svt}}$ even if attention is limited to indices j that (taken individually) permit reliable identification of $[j]_{\text{svt}}$.

Proposition 8B. *There is a set $X \subseteq N$ such that*

(i) *X is performable on $[\cdot]_{\text{svt}}$;*

(ii) *for all $j \in X$, $[j]_{\text{svt}}$ is indentifiable reliably; but*

(iii) *X is not performable on $[\cdot]_{\text{svt}}$ reliably.*

Proof. Recall the collection $\mathrm{RE}_{\mathrm{si}}$, introduced in Definition 4C.

We define total $h \in F^{\mathrm{rec}}$ such that for all $i \in N$, $h(i)$ is an index for the language that represents the function θ operating as follows. Given $j \in N$, θ computes $\phi_i(0)$, $\phi_i(1)$, ... until the smallest $x_0 \in N$ is found such that $\phi_i(x_0) = 1$. If there is no such x_0, or if $\phi_i(m)$ diverges for some $m < x_0$, then $\theta(j)$ is undefined. If such an x_0 is found, then

$$\begin{aligned}\theta(j) &= \phi_i(j), && \text{if } \phi_i(j) = \phi_{x_0}(j) \text{ and both computations converge;}\\ &= \text{undefined}, && \text{otherwise.}\end{aligned}$$

Observe that for all $i \in N$, if $W_i \in \mathrm{RE}_{\mathrm{si}}$, then $W_i = W_{h(i)} \in \mathrm{RE}_{\mathrm{svt}}$; whereas if $W_i \notin \mathrm{RE}_{\mathrm{si}}$, then $W_{h(i)} \notin \mathrm{RE}_{\mathrm{svt}}$.

Let total, one–one $d \in F^{\mathrm{rec}}$ be such that for all $i \in N$, $W_{d(i)} = \{h(i)\}$. Thus, for all $i \in N$, $[d(i)]_{\mathrm{svt}} = \{W_i\}$ if $W_i \in \mathrm{RE}_{\mathrm{si}}$; $= \varnothing$ otherwise. Let X of the proposition be $\{d(i) \mid i \in N\}$.

Clause (i) follows from Proposition 3B, and (ii) is easy to verify. Suppose for a contradiction that synthesizer S performs X on $[\cdot]_{\mathrm{svt}}$ reliably. Then, for all $i \in N$, $\lambda\sigma \cdot S(d(i), \sigma)$ is reliable and identifies W_i if $W_i \in \mathrm{RE}_{\mathrm{si}}$. Let $\theta \in F^{\mathrm{rec}}$ be defined as follows. Given $\sigma \in \mathrm{SEQ}$, θ puts out $\mathrm{lh}(\sigma)$ if there is no $i \in N$ such that $\langle i, 1\rangle \in \mathrm{rng}(\sigma)$. Otherwise, θ finds the least i with this property, and puts out $S(d(i), \sigma)$. It may be seen that θ is reliable and identifies $\mathrm{RE}_{\mathrm{si}}$. However, this contradicts Corollary 4.6.1B of OSW, due to Blum and Blum (1975). ∎

As a corollary to the foregoing proof, note that the set X of Proposition 8B can be chosen so that X is r.e. and for all $j \in X$, $\mathrm{card}(W_j) = 1$.

9. Memory Limitation

Inductive inference machines that do not store the data fed to them may be considered to have limited memory, and thus to conserve spatial resources in at least one sense.

Definition 9A (Wexler and Culicover, 1980). (i) $\theta \in F^{\mathrm{rec}}$ is *memory-limited* just in case for all σ, $\tau \in \mathrm{SEQ}$, if $\sigma(\mathrm{lh}(\sigma) - 1) = \tau(\mathrm{lh}(\tau) - 1)$ and $\theta(\sigma[\mathrm{lh}(\sigma) - 1]) = \theta(\tau[\mathrm{lh}(\tau) - 1])$, then $\theta(\sigma) = \theta(\tau)$.

(ii) $\mathbf{L} \subseteq \mathrm{RE}$ is *identifiable with limited memory* just in case some memory-limited $\theta \in F^{\mathrm{rec}}$ identifies $\mathbf{L}$.

In other words, θ is memory-limited just in case for all $\sigma \in \mathrm{SEQ}$, $\theta(\sigma)$ depends on no more than the last member of σ and θ's previous conjecture.

DEFINITION 9B. Let $X \subseteq N$, synthesizer S, and description function $\mathbf{D}$ be given.

(i) *S performs X on $\mathbf{D}$ with limited memory* just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$ and $\lambda\sigma \cdot S(i, \sigma)$ is memory-limited.

(ii) *X is performable on $\mathbf{D}$ with limited memory* just in case some synthesizer performs X on $\mathbf{D}$ with limited memory.

We now show that the requirement of limited memory can obstruct performability on $[\cdot]_{\mathrm{svt}}$ even if attention is limited to indices j that (taken individually) permit identification of $[j]_{\mathrm{svt}}$ with limited memory.

PROPOSITION 9A. *There is a set $X \subseteq N$ such that*

(i) *X is performable on $[\cdot]_{\mathrm{svt}}$;*

(ii) *for all $j \in X$, $[j]_{\mathrm{svt}}$ is identifiable with limited memory; but*

(iii) *X is not performable on $[\cdot]_{\mathrm{svt}}$ with limited memory.*

The following variant of Lemma 7A will be helpful.

LEMMA 9A. *Let $\theta \in F^{\mathrm{rec}}$ identify $L \in \mathrm{RE}$, and let $\sigma \in \mathrm{SEQ}$ be such that $\mathrm{rng}(\sigma) \subseteq L$. Then there is $\tau \in \mathrm{SEQ}$ such that* (a) $\mathrm{rng}(\tau) \subseteq L$, (b) $W_{\theta(\sigma \wedge \tau)} = L$, *and* (c) *for all $\delta \in \mathrm{SEQ}$, if $\mathrm{rng}(\delta) \subseteq L$ then $\theta(\sigma \wedge \tau \wedge \delta) = \theta(\sigma \wedge \tau)$.*

Proof. This is Corollary 2.1A of OSW. ▌

Proof of Proposition 9A. Given $S \subseteq N$, the set $\{\langle x, n_x \rangle \mid n_x = 0$ if $x \in S$, and $n_x = 1$ otherwise$\}$ is denoted C_S. We now demonstrate the existence of total $d \in F^{\mathrm{rec}}$ such that for all $i \in N$:

(1) if W_i is finite, then $[d(i)]_{\mathrm{svt}} = \{C_E \mid E \subseteq N$ finite$\}$; whereas,

(2) if W_i is infinite, then $[d(i)]_{\mathrm{svt}} = \{C_N\}$.

To show the existence of such a d, let total $f \in F^{\mathrm{rec}}$ be such that for all $i, p, y \in N$,

$$W_{f(\langle i,p,y \rangle)} = C_{D_y} \quad \text{if } W_{i,p} = W_i;$$
$$= C_{D_y} \cup \{\langle 0, 0 \rangle, \langle 0, 1 \rangle\} \quad \text{otherwise.}$$

(Recall that D_y is the yth finite set.) Observe that for all $i, p, y \in N$, $W_{f(\langle i,p,y \rangle)} \in \mathrm{RE}_{\mathrm{svt}}$ iff $W_{i,p} = W_i$.

Let total $g \in F^{\mathrm{rec}}$ be such that for all $i \in N$, $W_{g(i)} = \{\langle x, 0 \rangle \mid \mathrm{card}(W_i) \geqslant x\}$. Observe that for all $i \in N$, $W_{g(i)} = C_N$ iff W_i is infinite; otherwise, $W_{g(i)} \notin \mathrm{RE}_{\mathrm{svt}}$.

Let total $d \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{d(i)} = \{g(i)\} \cup \{f(\langle i, p, y\rangle) \mid p, y \in N\}$. If W_i is finite, then $W_i = W_{i,p}$ for some $p \in N$; so, $\{f(\langle i, p, y\rangle) \mid y \in N\}$ is an index set for $\{C_E \mid E \subseteq N \text{ finite}\}$ and $W_{g(i)} \notin \text{RE}_{\text{svt}}$. Thus, if W_i is finite, $[d(i)]_{\text{svt}} = \{C_E \mid E \subseteq N \text{ finite}\}$, verifying (1). If W_i is infinite, then $W_i \neq W_{i,p}$ for all $p \in N$. In this case, it is easy to see that $[d(i)]_{\text{svt}} = \{C_N\}$, verifying (2).

Let X of the proposition be $\{d(i) \mid i \in N\}$. Clauses (i) and (ii) are easy to verify. For a contradiction, let synthesizer S perform X on $[\cdot]_{\text{svt}}$ with limited memory. We say that σ is *N-like* if $\text{rng}(\sigma) \subseteq C_N$. We say that $q \in N$ is *out of* $\sigma \in \text{SEQ}$ just in case for all $n \in N$, $\langle q, n\rangle \notin \text{rng}(\sigma)$. Let us show that for all $i \in N$, if W_i is finite, then:

$$(\forall \sigma \in \text{SEQ})\ (\exists q \in N)\ (\sigma \text{ is } N\text{-like} \rightarrow .\ q \text{ is out of } \sigma \text{ and } S(d(i), \sigma) \neq S(d(i), \sigma \wedge \langle q, 0\rangle)). \tag{3}$$

For suppose otherwise, and let $i \in N$ be such that W_i is finite. Let $\sigma \in \text{SEQ}$ be such that σ is N-like and for all $q \in N$, if q is out of σ then $S(d(i), \sigma) = S(d(i), \sigma \wedge \langle q, 0\rangle)$. Let $A = \{x \in N \mid \langle x, 0\rangle \in \text{rng}(\sigma)\}$. By (1), $\lambda\tau \cdot S(d(i), \tau)$ identifies C_A. By Lemma 9A let $\tau \in \text{SEQ}$ be such that $\text{rng}(\tau) \subseteq C_A$, $W_{S(d(i), \sigma \wedge \tau)} = C_A$, and for all $\delta \in \text{SEQ}$, if $\text{rng}(\delta) \subseteq C_A$ then $S(d(i), \sigma \wedge \tau \wedge \delta) = S(d(i), \sigma \wedge \tau)$. Let q be out of $\text{rng}(\sigma \wedge \tau)$. Then, $S(d(i), \sigma) = S(d(i), \sigma \wedge \langle q, 0\rangle)$. So, by the memory-limitedness of $\lambda\sigma \cdot S(d(i), \sigma)$, $S(d(i), \sigma \wedge \langle q, 0\rangle \wedge \tau \wedge \delta) = S(d(i), \sigma \wedge \tau \wedge \delta)$ for all $\delta \in \text{SEQ}$. Now let t be a text for $C_A - \{\langle q, 1\rangle\}$, and let s be the text $\sigma \wedge \langle q, 0\rangle \wedge \tau \wedge t$. It may be seen that s is a text for $C_{A \cup \{q\}} \neq C_A$, and that $\lambda\sigma \cdot S(d(i), \sigma)$ converges on s to $S(d(i), \sigma \wedge \tau)$, i.e., to an index for $C_A \neq \text{rng}(s) = C_{A \cup \{q\}}$. Thus, $\lambda\sigma \cdot S(d(i), \sigma)$ does not identify $C_{A \cup \{q\}}$, contradicting (1) and the assumption that S performs X. This establishes (3).

On the other hand, for all $i \in N$, if W_i is infinite, then:

$$(\exists \sigma \in \text{SEQ})\ (\forall q \in N)\ (\sigma \text{ is } N\text{-like and } q \text{ is out of } \sigma\ . \rightarrow S(d(i), \sigma) = S(d(i), \sigma \wedge \langle q, 0\rangle)). \tag{4}$$

This claim follows directly from Lemma 7A along with the observation that if W_i is infinite, then $\lambda\sigma \cdot S(d(i), \sigma)$ identifies C_N.

Now observe that for all i, $q \in N$, $\sigma \in \text{SEQ}$, if σ is N-like then both $S(d(i).\,\sigma)$ and $S(d(i), \sigma \wedge \langle q, 0\rangle)$ are defined. For otherwise, $\lambda\tau \cdot S(d(i), \tau)$ fails to identify some $L \in [d(i)]_{\text{svt}}$. As a consequence, the matrix expressions in (3) and (4) represent decidable relations. Hence (3) and (4) exhibit $\{i \in N \mid W_i \text{ finite}\}$ as Π_2, contradicting its Σ_2-completeness (see Rogers, 1967, Section 14.8). ∎

10. Incomplete Text

Useful inductive inference machines should tolerate small deformations of the data they examine. In the present section we consider the synthesis of machines that succeed on texts with very small alterations, namely, the introduction of a single "gap."

Definition 10A. Let $L \in \mathrm{RE}_{\mathrm{svt}}$ and text t be given. t is an *incomplete text for* L just in case $\mathrm{rng}(t) = L - \{\langle n, m\rangle\}$ for some $n, m \in N$.

Since $\langle n, m\rangle \notin L$ is possible, we see that every text for $L \in \mathrm{RE}_{\mathrm{svt}}$ counts as an incomplete text for L. An incomplete text for $L \in \mathrm{RE}_{\mathrm{svt}}$ may be conceived as the result of removing every occurrence of at most one number $\langle n, m\rangle$ from some text for L.

Definition 10B. Let $\theta \in F^{\mathrm{rec}}$ and $L \in \mathrm{RE}_{\mathrm{svt}}$ be given.

(i) *θ identifies L with incomplete text* just in case for every incomplete text t for L, θ converges on t to an index for L.

(ii) $\mathbf{L} \subseteq \mathrm{RE}_{\mathrm{svt}}$ is *identifiable with incomplete text* just in case there is $\theta \in F^{\mathrm{rec}}$ such that for every $L \in \mathbf{L}$, θ identifies L with incomplete text.

It follows from the definition that if $\theta \in F^{\mathrm{rec}}$ identifies $\mathbf{L} \subseteq \mathrm{RE}_{\mathrm{svt}}$ with incomplete text, then θ identifies $\mathbf{L}$. Let $L, L' \in \mathrm{RE}_{\mathrm{svt}}$ be such that $\mathrm{card}(L - L') = 1$. Then $\{L, L'\}$ is not identifiable with incomplete text.

Definition 10C. Let $X \subseteq N$, synthesizer S, and description function $\mathbf{D}$ be given.

(i) *S performs X on $\mathbf{D}$ with incomplete text* just in case for all $i \in X$, $\lambda\sigma \cdot S(i, \sigma)$ identifies $\mathbf{D}(i)$ with incomplete text.

(ii) *X is performable on $\mathbf{D}$ with incomplete text* just in case some synthesizer S performs X on $\mathbf{D}$ with incomplete text.

The following proposition shows that in the context of SVT, incompletion has no genuine impact on performability.

Definition 10D. (i) $i \in N$ is *varied* just in case for all $j, k \in W_i$, if $j \neq k$ then $\mathrm{card}(W_j - W_k) \neq 1$.

(ii) $\mathrm{VAR} = \{i \in N \mid i \text{ is varied}\}$.

Proposition 10A. $\mathrm{SVT} \cap \mathrm{VAR}$ *is performable on* $[\cdot]_{\mathrm{svt}}$ *with incomplete text.*

Proof. This is an easy variant of the proof of Proposition 4A. ∎

In the more general context, the next result shows that the requirement embodied in Definition 10C can obstruct performability on $[\cdot]_{svt}$, even if attention is limited to indices j that (taken individually) permit identification of $[j]_{svt}$ with incomplete text.

PROPOSITION 10B. *There is a set $X \subseteq N$ such that*

(i) *X is performable on $[\cdot]_{svt}$;*

(ii) *for all $j \in X$, $[j]_{svt}$ is identifiable with incomplete text; but*

(iii) *X is not performable on $[\cdot]_{svt}$ with incomplete text.*

A lemma will be helpful. A synthesizer S is called *total* just in case S is defined on all of $N \times \text{SEQ}$.

LEMMA 10A. *Let $X \subseteq N$ and description function $\mathbf{D}$ be given. Suppose that X is performable on $\mathbf{D}$ with incomplete text. Then there is a total synthesizer S that performs X on $\mathbf{D}$ with incomplete text.*

Proof. The lemma is an elementary adaptation of Lemma 4.2.2B of OSW. ∎

Proof of Proposition 10B. Let total $h \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{h(i)}$ is enumerated as follows:

Stage 0. Enumerate $\langle 0, 8 \rangle$ into $W_{h(i)}$.

Stage $x+1$. Enumerate $\langle x+1, 0 \rangle$ into $W_{h(i)}$ iff $\text{card}(W_i) \geq x+1$; diverge if $\text{card}(W_i) < x+1$.

Thus, $W_{h(i)} \in \text{RE}_{svt}$ iff W_i is infinite.

Let total $g \in F^{\text{rec}}$ be such that for all $i, m \in N$, $W_{g(\langle i,m \rangle)}$ is enumerated as follows:

Stage 0. Enumerate $\langle 0, 9 \rangle$ into $W_{g(\langle i,m \rangle)}$. Then find the smallest $p \in N$ such that $\text{card}(W_{i,p}) = m$; diverge if no such p is found.

Stage $x+1$. Enumerate $\langle x+1, 0 \rangle$ into $W_{g\langle i,m \rangle}$ iff $W_{i,p} = W_{i,p+x+1}$. Diverge if $W_{i,p} \subset W_{i,p+x+1}$.

Thus, $W_{g(\langle i,m \rangle)} \in \text{RE}_{svt}$ iff $\text{card}(W_i) = m$.

Let total $d \in F^{\text{rec}}$ be such that for all $i \in N$, $W_{d(i)} = \{h(i)\} \cup \{g(\langle i, m \rangle) \mid m \in N\}$. Observe that for all $i \in N$, $W_{d(i)}$ contains indices for exactly one $L \in \text{RE}_{svt}$. Let X of the proposition be $\{d(i) \mid i \in N\}$. Clauses (i) and (ii) are easy to verify.

For a contradiction, suppose that X is performable on $[\cdot]_{svt}$ with incomplete text. By Lemma 10A let total synthesizer S perform X on $[\cdot]_{svt}$ with incomplete text. Let $i \in N$ be given, and let text $t = \langle 1, 0 \rangle$, $\langle 2, 0 \rangle$,

$\langle 3, 0\rangle$, ... Observe that t is an incomplete text for $W_{h(i)}$ iff W_i is infinite, and that t is an incomplete text for $W_{g(\langle i,m\rangle)}$ iff $\text{card}(W_i) = m$.

Now suppose that W_i infinite. Then $\lambda\sigma \cdot S(d(i), \sigma)$ identifies $W_{h(i)} \in \text{RE}_{\text{svt}}$ with incomplete text. So, $\lambda\sigma \cdot S(d(i), \sigma)$ converges on t to an index for $W_{h(i)}$. This shows for all $i \in N$, if W_i is infinite then:

$$(\exists n, x \in N)\ (\forall j \in N)\ (\langle 0, 8\rangle \in W_{S(d(i), t[n]), x} \text{ and } S(d(i), t[n]) = S(d(i), t[n+j])). \tag{5}$$

On the other hand, suppose that $\text{card}(W_i) = m$. Then $\lambda\sigma \cdot S(d(i), \sigma)$ identifies $W_{g(\langle i,m\rangle)}$ with incomplete text. So, $\lambda\sigma \cdot S(d(i), \sigma)$ converges on t to an index for $W_{g(\langle i,m\rangle)}$. Since $\langle 0, 8\rangle \notin W_{g(\langle i,m\rangle)}$, this shows for all $i \in N$, if W_i is finite then:

$$(\forall n, x \in N)(\exists j \in N)(\langle 0, 8\rangle \in W_{S(d(i), t[n]), x} \rightarrow S(d(i), t[n]) \neq S(d(i), t[n+j])). \tag{6}$$

Since S is total, (5) and (6) exhibit $\{i \in N \mid W_i \text{ infinite}\}$ as Σ_2, whereas it is Π_2-complete. ∎

11. Concluding Remarks

One extension of the foregoing study would be to taxonomize description functions according to their computational or informational transparency, i.e., in terms of their aptness for communicating with synthesizers. Two concepts that suggest themselves for further study in this regard may be defined as follows.

Definition 11A. Let description function $\mathbf{D}$ and $X \subseteq N$ be given.

(i) $\mathbf{D}$ is *normal on* X just in case for all $i \in X$, $\mathbf{D}(i) \subseteq [i]$.

(ii) $\mathbf{D}$ is *well behaved on* X just in case there is a total recursive g such that for all $i \in X$, $\mathbf{D}(i) = [g(i)]$.

It may be seen that $[\cdot]$ is both normal and well behaved on N whereas $[\cdot]_{\text{svt}}$ is normal on N but not well behaved on N. Indeed, it is easy to show that $[\cdot]_{\text{svt}}$ is not even well behaved on $\{i \mid \text{card}(W_i) = 1\}$. On the other hand, $[\cdot]_{\text{svt}}$ is well behaved on SVT and Proposition 4A can be generalized to say that X is performable on $[\cdot]_{\text{svt}}$ whenever $[\cdot]_{\text{svt}}$ is well behaved on X. It may also be remarked that the rather natural description function $\mathbf{D}(i) = \{W_j \mid j \in D_i\}$ is well behaved on N but not normal on N. Propositions 3B and 4E provide information about performability on this $\mathbf{D}$.

The informational and computational opacity of description functions which fail to be either normal or well behaved is illustrated by the following proposition, the proof of which is omitted.

DEFINITION 11B. (i) Let description function $[\cdot]_c$ be defined as follows: For all $i \in N$, $[i]_c = \{W_j \mid j \in W_i^c\}$.

(ii) $\mathrm{EXT}_c = \{i \in N \mid (\forall j, k \in W_i^c)\ (j \neq k \rightarrow W_j \neq W_k)\}$.

PROPOSITION 11A. *$\{i \in \mathrm{EXT}_c \mid \mathrm{card}(W_i^c) \leqslant 2\}$ is not performable on $[\cdot]_c$.*

This result should be compared to Proposition 3B.

Another area for further study concerns the nature of those design specifications for which results of the kind given in Sections 5–9 hold. The main propositions of these sections have a similar form, and it would be useful to extract a property shared by the five design specifications there treated from which the results of Sections 5–9 might be uniformly derived.

RECEIVED September 21, 1986; ACCEPTED September 1987

REFERENCES

ANGLUIN, D. (1980), Inductive inference of formal languages from positive data, *Inform. and Control* **45**, 117–135.

BLUM, L. AND BLUM, M. (1975), Toward a mathematical theory of inductive inference, *Inform. and Control* **28**, 125–155.

CASE, J. AND SMITH, J. (1983), Comparison of identification criteria for machine inductive inference, *Theoret. Comput. Sci.* **25**, 193–220.

GOLD, E. M. (1967), Language identification in the limit, *Inform. and Control* **10**, 447–474.

OSHERSON, D., STOB, M., AND WEINSTEIN, S. (1986), "Systems that Learn" MIT Press, Cambridge, MA.

WEXLER, K., AND CULICOVER, P. (1980), "Formal Principles of Language Acquisition," MIT Press, Cambridge, MA.